

Entry time statistics to different shrinking sets

Italo Cipriano

April 2016

Abstract

We consider ψ -mixing dynamical systems $(\mathcal{X}, T, \mathcal{B}_{\mathcal{X}}, \mu)$ and we find conditions on families of sets $\{\mathcal{U}_n \subset \mathcal{X} : n \in \mathbb{N}\}$ so that $\mu(\mathcal{U}_n)\tau_n$ tends in law to an exponential random variable, where τ_n is the entry time to \mathcal{U}_n .

1 Introduction

We develop further results about higher order entry times. That is the rate at which points enter to small sets. Consider a measure preserving dynamical system $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T)$ where μ is a finite, invariant and ergodic measure together with a sequence of Borel sets $\{\mathcal{U}_n\} = \{\mathcal{U}_n\}_{n \in \mathbb{N}}$ with $\mathcal{U}_n \subset \mathcal{X}, \mu(\mathcal{U}_n) > 0$ and for which the sequence $\{\mathcal{U}_n\}$ shrinks to a point. Under an appropriate mixing assumption, one can generally show that for

$$\tau_n(x) = \tau_{\mathcal{U}_n}(x) := \inf\{k \geq 1 : T^k(x) \in \mathcal{U}_n\},$$

the sequence of random variables $X_n := \mu(\mathcal{U}_n)\tau_n$, called (rescaled) *entry time* or *first hitting time*, converges in law to an exponential random variable. The first paper related to this result is [14] for continued fractions. More recently such convergence results have been obtained for examples in which \mathcal{U}_n are balls or cylinders shrinking to a point: for continuous time Markov chains see [4, 5], for expanding maps of the interval see [11–13], for general ϕ -mixing processes (consequently also for ψ -mixing processes) see [16], for α -mixing processes see [1, 3], for Axiom A diffeomorphisms see [20, 21], for Gibbs measures on shift spaces see [23], for unimodal maps see [7], for partially hyperbolic systems see [8, 15]. Some extensions of the classic Poisson limit theorem can be found in [24] and further related reviews in [2, 10, 17]. It is not clear however that the limit distribution of the sequence of random variables X_n is still valid if we remove the condition that $\{\mathcal{U}_n\}$ are shrinking to a single point.

In this paper we consider the problem of finding conditions on the pair $(\{\mathcal{U}_n\}, \mu)$ so that X_n converges in law to an exponential random variable when $\mu(\mathcal{U}_n) \rightarrow 0$ as $n \rightarrow \infty$, but the sequence $\{\mathcal{U}_n\}$ does not necessarily shrink to a single point (or finitely many points). The conditions we impose on \mathcal{U}_n and μ are similar to the used in [9, 18, 19]. In many cases these conditions are easy to verify, this allows us to find new systems satisfying exponential limit distribution.

In our main theorem we obtain the convergence of X_n to an exponential random variable for general families of sets $\{\mathcal{U}_n\}$ under some conditions that depend on the return time to \mathcal{U}_n defined by

$$\eta_n := \inf\{\tau_n(x) : x \in \mathcal{U}_n\}.$$

Indeed, let (\mathcal{X}, σ) be a topologically mixing subshift of finite type with

$$\mathcal{X} \subset \prod_{-\infty}^{\infty} \{1, \dots, a\} = \{1, \dots, a\}^{\mathbb{Z}}.$$

We consider the pair $(\{\mathcal{U}_n\}, \mu)$, where $\mathcal{U}_n \subset \mathcal{X}$ is in the sigma-algebra generated by $\prod_{-n}^{n-1} \{1, \dots, a\}$ and $\mu = \mu_{\text{Parry}}$ is the probability measure of maximum entropy or Parry measure.

Theorem 1.1 (Main Theorem). *The sequence of random variables X_n converges in law to an exponential random variable if*

- i. $\sigma^{-(n+1)}\mathcal{U}_{n+1} \subset \sigma^{-n}\mathcal{U}_n$ for every $n > 0$ and $n\mu(\mathcal{U}_{\lfloor \eta_n/2 \rfloor}) \rightarrow 0$ as $n \rightarrow \infty$,
- ii. the return times are given by $\eta_n = n + k(n) + 1$, where $k : \mathbb{N} \rightarrow \mathbb{N}$ is a non decreasing function and $n\mu(\mathcal{U}_{k(n)}) \rightarrow 0$ as $n \rightarrow \infty$, or
- iii. there exists a sequence $\{\mathcal{V}_n\}$ with $n\mu(\mathcal{V}_n) \rightarrow 0$ as $n \rightarrow \infty$, where $\mathcal{V}_n \supset \mathcal{U}_n$ is a set in the sigma-algebra generated by $\prod_{-n}^{-n+\lfloor \eta_n/2 \rfloor} \{1, \dots, a\}$ or $\prod_{n-\lfloor \eta_n/2 \rfloor}^{n-1} \{1, \dots, a\}$ for every $n \geq 1$.

An application of our method is to the study of Gibbs measures and sets which do not necessarily shrink to single points. Indeed, suppose that we have two subshifts of finite type \mathcal{X} and $\mathcal{Y} \subset \mathcal{X}$, and the family of sets $\{\mathcal{U}_n\}, \mathcal{U}_n \subset \mathcal{X}$ satisfies $\cap_{n=1}^{\infty} \mathcal{U}_n = \{xy : y \in \mathcal{Y}\}$, where x is a forbidden sequence in \mathcal{Y} . We prove that for μ a Gibbs measure of Hölder potential, and under suitable conditions, that depend on μ and the topological entropy of (\mathcal{X}, σ) and (\mathcal{Y}, σ) , the sequence X_n converges in law to an exponential random variable. This is closely related to the results in [9] but our proofs are motivated by [11]. Recently, a necessary and sufficient condition for X_n to converge in law to an exponential random variable when μ is an ergodic probability measure is established in Theorem 1 in [25], however, to check this condition requires not straightforward estimations.

2 Background

We will formally introduce the definition of subshift of finite type. Let A denote an irreducible and aperiodic $a \times a$ matrix of zeros and ones ($a \geq 2$), i.e. there exists $d \in \mathbb{N}$ for which $A^d > 0$ (all coordinates of A^d are strictly positive). We call the matrix A transition matrix. We define the subshift of finite type $\mathcal{X} = \mathcal{X}_A \subset \{1, \dots, a\}^{\mathbb{Z}}$ such that

$$\mathcal{X} := \{(x_n)_{n=-\infty}^{\infty} : A(x_n, x_{n+1}) = 1 \text{ for all } n \in \mathbb{Z}\}.$$

On \mathcal{X} , the shift $\sigma : \mathcal{X} \rightarrow \mathcal{X}$ is defined by $\sigma(x)_n = x_{n+1}$ for all $n \in \mathbb{Z}$. For $x \in \mathcal{X}$ and $n, m \in \mathbb{Z}, m > n$ we define the cylinder

$$[x]_n^m := \{y \in \mathcal{X} : y_i = x_i, \text{ for } i \in \{n, \dots, m-1\}\}$$

and also denote $x_{[n,m)} := x_n x_{n+1} \dots x_{m-1} = (x_k)_{k=n}^{m-1}$, which corresponds to the concatenation of $m - n$ elements in $\{1, \dots, a\}$. We denote by ξ_n^m the set of all the cylinders $[x]_n^m$ with $x \in \mathcal{X}$. In the particular, when $n = 0$ we denote $[x]_n^m =: [x]_m$ and $\xi_n^m =: \xi_m$. Denote by \mathcal{M}_σ the space of σ invariant probability measure on \mathcal{X} , that is the space of probability measures μ on $\mathcal{B}_\mathcal{X}$ such that $\mu(\mathcal{A}) = \mu(\sigma^{-1}\mathcal{A})$ for every measurable set \mathcal{A} .

For $\theta \in (0, 1)$, we consider the metric $d_\theta(x, y) = \theta^m$ where $m = \inf\{n \geq 0 : x_n \neq y_n \text{ or } x_{-n} \neq y_{-n}\}$ and $d(x, x) = 0$ for every $x \in \mathcal{X}$, we have in particular that (\mathcal{X}, d_θ) is a complete metric space. We say that $f : \mathcal{X} \rightarrow \mathbb{R}$ is continuous if it is continuous with respect to d_θ . Given $f : \mathcal{X} \rightarrow \mathbb{R}$ continuous and $m \in \mathbb{N}$ define

$$V_m(f) := \sup\{|f(x) - f(y)| : x, y \in \mathcal{X} \text{ and } x_i = y_i \ \forall i \in \{-m, \dots, m\}\},$$

and the Lipschitz semi-norm

$$|f|_\theta := \sup \left\{ \frac{V_m(f)}{\theta^m} : m \in \mathbb{N} \right\}.$$

Since constant functions all have Lipschitz semi-norm equal to zero, one needs to define the norm on the space of Lipschitz functions by

$$\|f\|_\theta := |f|_\theta + \|f\|,$$

where $\|f\| := \sup_{x \in \mathcal{X}} \{|f(x)|\}$. The space of continuous functions with finite Lipschitz norm is called the space of Lipschitz functions (or θ -Lipschitz functions). Recall that a continuous function is α -Hölder for d_θ if and only if it is Lipschitz for d_{θ^α} .

Definition 2.1. *Given a subshift of finite type (\mathcal{X}, σ) with incidence matrix A of size $a \times a$, the topological entropy h of \mathcal{X} is defined as the logarithm of the first eigenvalue of A , i.e., for $\lambda := \max\{|\gamma| : Av = \gamma v, \gamma \in \mathbb{C}, v \in \mathbb{C}^a\}$, $h := \log \lambda$. On the space of invariant probability measures \mathcal{M}_σ there exists one and only one measure μ (see [22]) such that*

$$h = - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mathcal{W} \in \xi_n} \mu(\mathcal{W}) \log \mu(\mathcal{W}). \quad (1)$$

The measure $\mu \in \mathcal{M}_\sigma$ that achieves the supremum in (1) is the called measure of maximum entropy or Parry measure, that we denote by μ_{Parry} .

Suppose that we have a subshift of finite type (\mathcal{X}, σ) and an invariant probability measure μ on $\mathcal{B}_\mathcal{X}$. We will need the following mixing condition on the measure μ .

Definition 2.2. *The measure preserving dynamical system $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mu, \sigma)$ is ψ -mixing if for \mathcal{U} in the σ -algebra generated by $\{1, \dots, a\}^n$ and \mathcal{V} in the σ -algebra generated by $\{1, \dots, a\}^*$ $:= \cup_{n \in \mathbb{N}} \{1, \dots, a\}^n$ as the ‘gap’ $\Delta \rightarrow \infty$:*

$$\sup_n \sup_{\mathcal{U}, \mathcal{V}} \left| \frac{\mu(\mathcal{U} \cap \sigma^{-\Delta-n}\mathcal{V})}{\mu(\mathcal{U})\mu(\mathcal{V})} - 1 \right| = \psi_\Delta \rightarrow 0.$$

Let us recall some basic definitions on probability theory that we will use.

Definition 2.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, i.e. Ω is a set, \mathcal{F} a sigma-algebra on Ω and \mathbb{P} is a probability measure on (Ω, \mathbb{F}) . The expectation with respect to \mathbb{P} of a measurable function $f : \Omega \rightarrow \mathbb{R}$ is defined by

$$\mathbb{E}(f) = \mathbb{E}_{\mathbb{P}}(f) := \int_{\Omega} f d\mathbb{P}.$$

Definition 2.4. A random variable $X : \Omega \rightarrow \mathbb{R}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be an exponential random variable with parameter λ if it has cumulative distribution

$$F_X(t) := \mathbb{P}\{w \in \Omega : X(w) \leq t\} = \begin{cases} 1 - e^{-\lambda t} & \text{if } t \geq 0, \\ 0 & \text{if } t < 0. \end{cases}$$

Definition 2.5. A sequence of random variables $\{X_n\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to converge in law to an exponential random variable of parameter 1 if for every $t \in \mathbb{R}$, its cumulative distribution converges to the cumulative distribution $F_X(t)$, of an exponential random variable X of parameter $\lambda = 1$. In other words,

$$\lim_{n \rightarrow \infty} |\mathbb{P}\{w \in \Omega : X_n(w) \geq t\} - e^{-t}| = 0 \text{ for every } t > 0. \quad (2)$$

We introduce the definition of some particular dynamical systems together with a sequence of sets whose measure converges to zero and a sequence of random variables built from the entry times to these sets.

Definition 2.6 (M -systems). We define a M -system as any system

$$(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T, \{\mathcal{U}_n\}, \{X_n\})$$

where

- i. the system $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T)$ is a measure preserving dynamical system and μ is an ergodic probability measure;
- ii. the sequence $\{\mathcal{U}_n\}$ is a sequence of Borel sets $\mathcal{U}_n \subset \mathcal{X}$ such that $\mu(\mathcal{U}_n) > 0$ and $\mu(\mathcal{U}_n) \rightarrow 0$; and
- iii. the sequence $\{X_n\}$ is a sequence of random variables $X_n : \mathcal{X} \rightarrow \mathbb{R}$ defined by $X_n(x) := \mu(\mathcal{U}_n) \tau_n(x)$.

A particularly useful M -system for us is the case that the measure is ψ -mixing.

Definition 2.7 (M_a -systems). We define a M_a -system as any M -system

$$(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma, \{\mathcal{U}_n\}, \{X_n\})$$

where (\mathcal{X}, σ) is a topologically mixing subshift of finite type with $\mathcal{X} \subset \{1, \dots, a\}^{\mathbb{Z}}$ for some integer $a > 1$ and the measure preserving dynamical system $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma)$ is ψ -mixing where the sequence $\{\psi_{\Delta}\}_{\Delta \in \mathbb{N}}$ is bounded.

We investigate the convergence in law of X_n in a M_a -system. A useful trick that we learned from [11] is to consider

$$\lim_{n \rightarrow \infty} \left| \mu \{x \in \mathcal{X} : \tau_n(x) > \lfloor t/\mu(\mathcal{U}_n) \rfloor\} - (1 - \mu(\mathcal{U}_n))^{\lfloor t/\mu(\mathcal{U}_n) \rfloor} \right| = 0 \text{ for every } t > 0, \quad (3)$$

that is equivalent to (2).

3 Intermediate results and examples

We present an important proposition, indeed a few improvements of it will prove Theorem 1.1.

Proposition 3.1 (Main proposition). *Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma, \{\mathcal{U}_n\}, \{X_n\})$ be a M_a -system where \mathcal{U}_n is in the sigma algebra generated by $\prod_{i=0}^{n-1} \{1, \dots, a\}$ for every $n \geq 1$. If*

$$\eta_n = n \quad (4)$$

and

$$n\mu(\mathcal{U}_n) \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (5)$$

Then the sequence of random variables X_n converges in law to an exponential random variable of parameter 1.

Notice that the condition $\eta_n = n$ is equivalent to $\eta_n \geq n$, because of the definition of the sequence of sets $\{\mathcal{U}_n\}$. As an application of the proposition we can give the following example, where $\mathcal{A} \sqcup \mathcal{B}$ denotes the disjoint union of the sets \mathcal{A} and \mathcal{B} .

Example 3.2. *Suppose that $\mathcal{X} = \{0, 1, 2\}^{\mathbb{Z}}$ and*

$$\mathcal{U}_n = \sqcup_{x_1, \dots, x_{n-1} \in \{1, 2\}} [0, x_1, x_2, \dots, x_{n-1}]_n.$$

Let μ be a Bernoulli probability measure on \mathcal{X} defined by a probability vector $(p_0, p_1, p_2) \in (0, 1)^3$. Then $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma, \{\mathcal{U}_n\}, \{X_n\})$ is a M_a -system and (4), (5) are satisfied, therefore Proposition 3.1 applies.

Another application is to Gibbs measures.

Definition 3.3 (Gibbs measures). *Given a subshift of finite type \mathcal{X} , we say that a probability measure μ on $\mathcal{B}_{\mathcal{X}}$ is a Gibbs measure (or an equilibrium state) of Hölder potential $\phi : \mathcal{X} \rightarrow \mathbb{R}$ if there is $c_1, c_2 > 0$ and $P \geq 0$ such that*

$$c_1 \leq \frac{\mu([x]_m)}{\exp(-Pm + S_m^\sigma \phi(x))} \leq c_2$$

for every $x \in \mathcal{X}$ and $m \geq 0$, where $S_m^\sigma \phi(x) := \sum_{k=0}^{m-1} \phi(\sigma^k x)$.

Remark 3.4. *Gibbs measures of Hölder potential are ψ -mixing, see [6].*

Before giving an example in the case of Gibbs measures let us mention that Proposition 3.1 gives a condition that depends on the pressure.

Remark 3.5. *Let (\mathcal{X}, σ) be a subshift of finite type and $\mathcal{U}_n = [x^1]_n \sqcup \dots \sqcup [x^{m_n}]_n$, where $x^1, \dots, x^{m_n} \in \mathcal{X}$, for every $n \geq 1$. If μ is a Gibbs measure on \mathcal{X} of Hölder potential $\phi : \mathcal{X} \rightarrow \mathbb{R}$, then*

$$0 \leq \mu(\mathcal{U}_n) \leq m_n \exp(-Pn + \|S_n^\sigma \phi\|), \quad (6)$$

for some constants $c > 0$. In particular, we deduce that the hypothesis

$$nm_n \exp(-Pn + \|S_n^\sigma \phi\|) \rightarrow 0 \text{ as } n \rightarrow \infty$$

is enough in order to satisfy the hypothesis (5).

Example 3.6. Let A be an irreducible and aperiodic matrix with entries 0 and 1 of size $a \times a$ with $a > 2$ and let B be a submatrix of A of size $b \times b$ with $b \in \{2, \dots, a-1\}$. Denote by λ_A and λ_B the Perron eigenvalues of A and B , respectively. Consider the subshift of finite type $\mathcal{X}_A \subset \{1, \dots, a\}^{\mathbb{Z}}$ and $\mathcal{X}_B \subset \mathcal{X}_A$. Suppose without loss of generality that $\mathcal{X}_B \subset \{1, \dots, b\}^{\mathbb{Z}}$. Define for $n \in \mathbb{N}$,

$$\mathcal{U}_n = \sqcup_{x_1, \dots, x_{n-1} \in \{1, 2, \dots, b\}} [ax_1 \cdots x_{n-1}]_n,$$

then the sequence of random variables $X_n := \mu(\mathcal{U}_n)\tau_n$ converges in law to an exponential random variable of parameter 1 for any equilibrium state with Hölder potential ϕ such that $3\|\phi\| < \lambda_A - \lambda_B$. As (5) is clearly satisfied, then Proposition 3.1 applies.

We extend Proposition 3.1 in order to allow short return times. In this case we do need to care about considering shrinking sets, unlike in previous construction.

Definition 3.7. We say that a sequence of sets $\{\mathcal{U}_n\}$ is a sequence of shrinking sets if $\mathcal{U}_n \supset \mathcal{U}_{n+1}$ for every $n \in \mathbb{N}$.

We can now state our first corollary.

Corollary 3.8 (First corollary). Suppose that $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma, \{\mathcal{U}_n\}, \{X_n\})$ is a M_a -system where \mathcal{U}_n is a Borel set in the sigma algebra generated by $\prod_0^{n-1} \{1, \dots, a\}$ for every $n \geq 1$. If the sequence of sets $\{\mathcal{U}_n\}$ is shrinking and it satisfies

$$n\mu(\mathcal{U}_{\lfloor \eta_n/2 \rfloor}) \rightarrow 0 \text{ as } n \rightarrow \infty, \quad (7)$$

then the sequence of random variables X_n converges in law to an exponential random variable of parameter 1.

Example 3.9. Let us choose $x \in \mathcal{X}$ and consider $\{\mathcal{U}_n\} = \{[x]_n : n \in \mathbb{N}\}$. If $n\mu([x]_{\lfloor \eta_n/2 \rfloor}) \rightarrow 0$, then the sequence of random variables $X_n := \mu([x]_n)\tau_n$ converges in law to an exponential random variable of parameter 1.

Example 3.10. In particular, for the measure of maximum entropy μ_{Parry} . If $\{\mathcal{U}_n\}$ is a shrinking sequence and $n\mu_{\text{Parry}}(\mathcal{U}_{\lfloor \eta_n/2 \rfloor}) \rightarrow 0$, then $X_n := \mu_{\text{Parry}}(\mathcal{U}_n)\tau_n$ converges in law to an exponential random variable of parameter 1. This result is not sharp, but to have certain control on η_n is necessary.

Example 3.11. Suppose that (\mathcal{X}, σ) is the full shift in two symbols, i.e. its transition matrix A is a two by two matrix with 1 in each coordinate, and suppose that μ is the uniform probability measure, i.e. $\mu([x]_n) = 2^{-n}$ for every $x \in \mathcal{X}$ and $n \in \mathbb{N}$. Then $X_n := \mu(\mathcal{U}_n)\tau_n$ converges in law to an exponential random variable of parameter 1 for any sequence of sets $\{\mathcal{U}_n\}$ such that $\eta_n = \log_2(n) + k_n$, where $\{k_n\} \subset \mathbb{N}$ is a divergent sequence.

4 Proofs

We have separated the proofs of our results into three subsections. In the first we introduce the notation that will be used along the section and we prove Proposition 3.1. In the second we prove Corollary 3.8. Finally, in the third we combine what we did in the first two subsections in order to prove Theorem 1.1.

4.1 Main proposition

The goal of this subsection is to prove Proposition 3.1. In what follows we suppose that $(\mathcal{X}, \sigma, \mu)$ is a subshift of finite type where μ is a probability measure on $\mathcal{B}_{\mathcal{X}}$ and that we have a sequence $\{\mathcal{U}_n\}$ of Borel sets $\mathcal{U}_n \subset \mathcal{X}$ such that $\mu(\mathcal{U}_n) \rightarrow 0$. Instead of (2) we consider the equivalent condition (3). We prove the assertion (3) in several steps. The first is to replace (3) by a limit involving the sum of N terms.

Lemma 4.1. *For $n \in \mathbb{N}$, $\epsilon_n = \mu(\mathcal{U}_n)$ and $N = \lceil t/\epsilon_n \rceil$ with $t > 0$ we have that*

$$\mu\{\tau_n > N\} - (1 - \epsilon_n)^N = \sum_{q=0}^{N-1} (1 - \epsilon_n)^{N-q-1} (\mu\{\tau_n > q+1\} - (1 - \epsilon_n)\mu\{\tau_n > q\}). \quad (8)$$

The proof of Lemma 4.1 consists in noticing that most of the term in the sum of the right hand side of the equation (8) cancel when summed with other term, the unique terms that do not cancel are indeed the ones on the left hand side of the equation.

For what follows we require additionally that μ is an invariant probability measure on $\mathcal{B}_{\mathcal{X}}$. For $p, q \in \mathbb{N}_0$ we use the notation $\mu\{\tau_p > q\}$ instead of $\mu\{x \in \mathcal{X} : \tau_p(x) > q\}$ and for every $n \in \mathbb{N}$ we define $\epsilon_n := \mu(\mathcal{U}_n)$ and $N := \lceil t/\epsilon_n \rceil$, where $t > 0$.

For $q \in \{0, \dots, N-1\}$ define

$$p_q(n) := (1 - \epsilon_n)^{N-q-1} (\mu\{\tau_n > q+1\} - (1 - \epsilon_n)\mu\{\tau_n > q\}),$$

$S_1(N) := \sum_{q=0}^{N-1} p_q(n)$ and $S_2(N) := \sum_{q=n}^{N-1} p_q(n)$. To obtain (3) we will show with the help of a few lemmas that

$$\mu\{\tau_n > N\} - (1 - \epsilon_n)^N = S_1(N) + S_2(N) \rightarrow 0 \text{ as } n \text{ tends to infinity.}$$

The second step of our proof is to bound the term $S_1(N)$. We have the following lemma.

Lemma 4.2. *For all $n \in \mathbb{N}$, we have that $S_1(N) \leq n\epsilon_n$.*

The proof is a direct consequence of a useful identity in the next lemma that requires the measure μ to be invariant. We denote by \mathbb{E} the expectation with respect to μ .

Lemma 4.3. *For all $n \in \mathbb{N}$, we have that for all $q \in \{0, \dots, N-1\}$*

$$\mu\{\tau_n > q+1\} - (1 - \epsilon_n)\mu\{\tau_n > q\} = \epsilon_n\mu\{\tau_n > q\} - \mu\{x \in \mathcal{U}_n : \tau_n(x) > q\}.$$

Proof. The result is direct from the definitions of the sets $\{\tau_n > q+1\}$ and $\{\tau_n > q\}$.

Indeed, we can write the following identities:

$$\begin{aligned}
& \mu\{\tau_n > q+1\} - (1 - \epsilon_n)\mu\{\tau_n > q\} \\
&= \mathbb{E} \left(\prod_{i=1}^{q+1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - (1 - \epsilon_n) \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\
&= \mathbb{E} \left(\prod_{i=1}^{q+1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - (1 - \epsilon_n) \mathbb{E} \left((\mathbb{1}_{\mathcal{U}_n^c} + \mathbb{1}_{\mathcal{U}_n}) \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\
&= \mathbb{E} \left(\prod_{i=1}^{q+1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - (1 - \epsilon_n) \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n^c} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - (1 - \epsilon_n) \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\
&= \mathbb{E} \left(\prod_{i=1}^{q+1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n^c} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\
&\quad + \epsilon_n \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n^c} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - (1 - \epsilon_n) \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\
&= \epsilon_n \mathbb{E} \left((1 - \mathbb{1}_{\mathcal{U}_n}) \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - (1 - \epsilon_n) \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\
&= \epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right).
\end{aligned}$$

Therefore we have

$$\begin{aligned}
& \mu\{\tau_n > q+1\} - (1 - \epsilon_n)\mu\{\tau_n > q\} \\
&= \epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right), \tag{9}
\end{aligned}$$

and this is enough to conclude the proof, because

$$\epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) = \epsilon_n \mu\{\tau_n > q\} - \mu\{x \in \mathcal{U}_n : \tau_n(x) > q\}.$$

□

In the third step, and most difficult one we bound $S_2(N)$. This step requires additionally that μ is ψ -mixing. Or equivalently, that $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma, \{\mathcal{U}_n\}, \{X_n\})$ is a M_a -system. In order to obtain an upper bound for $S_2(N)$ we will state a lemma with some intermediate bounds.

Lemma 4.4. *For all $n, k \in \mathbb{N}$, we have*

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \mathbb{1}_{\mathcal{U}_n} \circ \sigma^{n+k}) \leq \epsilon_n^2(1 + \psi_k), \tag{10}$$

and for $q \in \{2n+1, \dots, N-1\}$ we have

$$\left| \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \leq \epsilon_n^2 \left(n + \sum_{i=0}^{n-1} \psi_i \right), \tag{11}$$

$$\left| \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \leq n\epsilon_n \quad (12)$$

and

$$\left| \epsilon_n \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \leq \epsilon_n \psi_n. \quad (13)$$

Proof. Inequality (10) requires the ψ -mixing condition and inequality (11) is a consequence of (10).

Proof of (10) We can use the ψ -mixing condition to conclude that

$$\left| \mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \mathbb{1}_{\mathcal{U}_n} \circ \sigma^{n+k}) - \epsilon_n^2 \right| = \left| \mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \mathbb{1}_{\mathcal{U}_n} \circ \sigma^{n+k}) - \mu(\mathcal{U}_n)^2 \right| \leq \epsilon_n^2 \psi_k.$$

Then trivially

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \mathbb{1}_{\mathcal{U}_n} \circ \sigma^{n+k}) \leq \epsilon_n^2 \psi_k + \epsilon_n^2,$$

as required.

Proof of (11) Let $n, k \in \mathbb{N}$ and $q \in \{2n+1, \dots, N-1\}$. We have

$$\begin{aligned} & \left| \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &= \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \cdot \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \cdot \left(1 - \prod_{i=n}^{2n-1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right) \\ &\leq \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \cdot \left(1 - \prod_{i=n}^{2n-1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right) \\ &\leq \sum_{i=n}^{2n-1} \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i \right) \leq \epsilon_n^2 \sum_{i=0}^{n-1} \psi_i + n\epsilon_n^2, \end{aligned}$$

where we have used (10) in the last inequality.

Proof of (12) Let $n, k \in \mathbb{N}$ and $q \in \{2n+1, \dots, N-1\}$. We have

$$\begin{aligned} & \left| \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &= \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \cdot \left(1 - \prod_{i=n}^{2n-1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right) \\ &\leq \mathbb{E} \left(1 - \prod_{i=n}^{2n-1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \leq \sum_{i=n}^{2n-1} \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \circ \sigma^i \right) = n\epsilon_n. \end{aligned}$$

Proof of (13) Let $n, k \in \mathbb{N}$ and $q \in \{2n+1, \dots, N-1\}$. We have

$$\begin{aligned} & \left| \epsilon_n \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\leq \psi_n \epsilon_n \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \leq \psi_n \epsilon_n, \end{aligned}$$

because of the condition of ψ -mixing. □

We can now bound $S_2(N)$.

Lemma 4.5. *For all $n \in \mathbb{N}$, we have that*

$$S_2(N) \leq 4(n+1)(t+1)\epsilon_n + (t+1)\psi_n + N\epsilon_n^2 \sum_{i=0}^{n-1} \psi_i.$$

Proof. Recall that $S_2(N) = \sum_{q=n}^{N-1} p_q(n)$. From (9) we have that

$$\begin{aligned} p_q(n) &= (1 - \epsilon_n)^{N-q-1} \left(\epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right) \\ &\leq \epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \end{aligned}$$

for every $q \in \{n, \dots, N-1\}$. For a fixed q we bound $p_q(n)$ by the sum of two terms:

$$\begin{aligned} &\epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\ &\leq \left| \epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \epsilon_n \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\quad + \left| \epsilon_n \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| =: S_{21}(N, q) + S_{22}(N, q). \end{aligned}$$

Notice that we have used (4) to obtain that

$$\mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) = \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right).$$

Our goal now is to bound $S_{21}(N, q)$ and $S_{22}(N, q)$. For the first term we have the simple inequality $S_{21}(N, q) \leq n\epsilon_n^2$, because

$$\begin{aligned} S_{21}(N, q) &\leq \epsilon_n \mathbb{E} \left(\left| \prod_{i=1}^{n-1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i - 1 \right| \cdot \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \\ &\leq \epsilon_n \mathbb{E} \left(1 - \prod_{i=1}^{n-1} \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \leq \epsilon_n \mathbb{E} \left(\sum_{i=1}^{n-1} \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i \right) = n\epsilon_n^2. \end{aligned}$$

To bound $S_{22}(N, q)$ we use Lemma 4.4. Suppose that $q > 2n$, then

$$\begin{aligned} S_{22}(N, q) &= \left| \epsilon_n \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\leq \left| \epsilon_n \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\quad + \epsilon_n \left| \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\quad + \left| \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=2n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &=: I_1(N, q) + I_2(N, q) + I_3(N, q). \end{aligned}$$

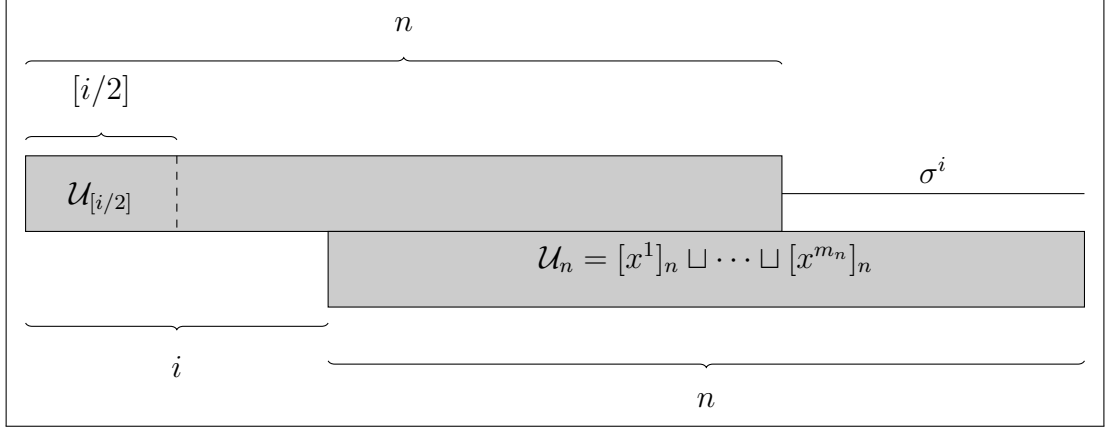


Figure 1: Proof of the inequality (14).

We have that $I_1(N, q) \leq \epsilon_n \psi_n$ by (13), $I_2(N, q) \leq n\epsilon_n^2$ by (12) and $I_3(N, q) \leq \epsilon_n^2 \sum_{i=0}^{n-1} \psi_i + n\epsilon_n^2$ by (11). In the case $n \leq q \leq 2n$ we can use that $S_{22}(N, q) \leq \epsilon_n$. Finally,

$$\begin{aligned}
S_2(N) &\leq \sum_{q=n}^N (S_{21}(N, q) + S_{22}(N, q)) \\
&= \sum_{q=n}^N S_{21}(N, q) + \sum_{q=n}^{2n} S_{22}(N, q) + \sum_{q=2n+1}^N S_{22}(N, q) \\
&\leq Nn\epsilon_n^2 + (n+1)\epsilon_n + N(I_1(N, q) + I_2(N, q) + I_3(N, q)) \\
&\leq \epsilon_n(n(t+1) + (n+1)) + (t+1)\psi_n + 2(t+1)n\epsilon_n + N\epsilon_n^2 \sum_{i=0}^{n-1} \psi_i,
\end{aligned}$$

which concludes the proof. \square

4.2 First corollary

The proof of the corollary is very similar to the one of Proposition 3.1, but we need to modify some details. To complete the proof we require the following lemma, where $n \in \mathbb{N}$ and $q > n$.

Lemma 4.6. *If $\eta_n = n$ then*

$$\left| \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \prod_{i=\eta_n}^q \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| = 0$$

and if $\eta_n < n$ then

$$\left| \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \prod_{i=\eta_n}^q \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \leq n\epsilon_n \epsilon_{\lfloor \eta_n/2 \rfloor} (1 + \psi_{\lfloor \eta_n/2 \rfloor}).$$

Proof. The case $\eta_n = n$ is trivial, so suppose that $\eta_n < n$. It is clear that

$$\left| \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \prod_{i=\eta_n}^q \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \leq \sum_{i=\eta_n}^{n-1} \mathbb{E} \left(\mathbf{1}_{\mathcal{U}_n} \cdot \mathbf{1}_{\mathcal{U}_n^c} \circ \sigma^i \right).$$

In Figure 1 we have chosen $i \in \{\eta_n, \dots, n-1\}$ and we represent the set $\mathcal{U}_n = [x^1]_n \sqcup \dots \sqcup [x^{m_n}]_n$ for some $x^1, \dots, x^{m_n} \in \mathcal{X}$, $m_n \in \mathbb{N}$ and the set $\sigma^i \mathcal{U}_n$. We have also draw a representation of the set $\mathcal{U}_{[i/2]}$. We can see that the action of the shift moved the rectangle at the bottom to the left, the result is the rectangle that we draw at the top. It is clear that

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) \leq \mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_{[i/2]}} \circ \sigma^i) = \mu(\mathcal{U}_n \cap \sigma^{-i} \mathcal{U}_{[i/2]}).$$

From the ψ -mixing condition of the measure μ we obtain that

$$|\mu(\mathcal{U}_n \cap \sigma^{-i} \mathcal{U}_{[i/2]}) - \epsilon_n \epsilon_{[i/2]}| \leq \psi_{[i/2]} \epsilon_{[i/2]} \epsilon_n,$$

and therefore

$$\begin{aligned} \sum_{i=\eta_n}^{n-1} \mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) &\leq \sum_{i=\eta_n}^{n-1} (\psi_{[i/2]} \epsilon_{[i/2]} \epsilon_n + \epsilon_n \epsilon_{[i/2]}) \\ &\leq n \epsilon_n \epsilon_{\lfloor \eta_n/2 \rfloor} (1 + \psi_{\lfloor \eta_n/2 \rfloor}). \end{aligned} \tag{14}$$

□

The proof of Corollary 3.8 comes from the observation that for $q \in \{n, \dots, N-1\}$ we have

$$\begin{aligned} &\left| \epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\leq \left| \epsilon_n \mathbb{E} \left(\prod_{i=1}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \epsilon_n \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\quad + \left| \epsilon_n \mathbb{E} \left(\prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &\quad + \left| \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=\eta_n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \\ &:= S_{21}(N, q) + S_{22}(N, q) + S_{\text{extra}}(N, q). \end{aligned}$$

We can use Lemma 4.6 to bound $S_{\text{extra}}(N, q)$ and the proof of Corollary 3.8 follows directly from the proof of Proposition 3.1.

4.3 Theorem 1.1

We can use the same notation and definitions used in the previous proofs and write

$$\mu(\tau_n > N) - (1 - \epsilon_n)^N = S_1(N) + S_2(N),$$

where

$$\begin{aligned} S_2(N) &\leq \sum_{q=n}^{N-1} S_{21}(N, q) + S_{22}(N, q) + S_{\text{extra}}(N, q) \text{ and} \\ S_{\text{extra}}(N, q) &:= \left| \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) - \mathbb{E} \left(\mathbb{1}_{\mathcal{U}_n} \prod_{i=\eta_n}^q \mathbb{1}_{\mathcal{U}_n^c} \circ \sigma^i \right) \right| \leq \sum_{i=\eta_n}^{2n-1} \mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i). \end{aligned}$$

Notice that the case i. is exactly the condition required in Corollary 3.8. Therefore, we only need to prove the theorem in the cases ii. and iii. The unique difference in these cases and the one of First corollary is that we need to find a new bound for $\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i)$ when $i \in \{\eta_n, \dots, 2n\}$.

Lemma 4.7. *If $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, \sigma, \{\mathcal{U}_n\}, \{X_n\})$ is a M_a -system where \mathcal{U}_n is in the sigma algebra generated by $\prod_{-n}^{n-1} \{1, \dots, a\}$ for every $n \geq 1$. Then, for every $n \geq 1$*

$$\begin{aligned} & \mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) \\ & \leq \begin{cases} \mu(\mathcal{U}_n)\mu(\mathcal{U}_k) + \psi_{\Delta}\epsilon_k\epsilon_n & \text{if } \eta_n = n + k + \Delta \text{ with } k + \Delta \geq 1, \\ \mu(\mathcal{V}_n)\mu(\mathcal{U}_n) + \psi_{\lfloor \eta_n/2 \rfloor}\mu(\mathcal{V}_n)\epsilon_n & \text{if } \eta_n < n + 1, \end{cases} \end{aligned}$$

for all $i \geq \eta_n$.

Proof. Let us fix $n \in \mathbb{N}$. We have two cases: $\eta_n = n + k + \Delta$ with $k + \Delta \geq 1$ or $\eta_n < n + 1$. Suppose first that $\eta_n = n + k + \Delta$ with $k + \Delta \geq 1$ and that $i \in \{\eta_n, \dots, 2n\}$, then

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) \leq \mathbb{E}(\mathbb{1}_{\mathcal{U}_k} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) = \mu(\mathcal{U}_k \cap \sigma^{-i}\mathcal{U}_n)$$

and

$$|\mu(\mathcal{U}_k \cap \sigma^{-i}\mathcal{U}_n) - \mu(\mathcal{U}_k)\mu(\mathcal{U}_n)| \leq \psi_{\Delta}\epsilon_k\epsilon_n. \quad (15)$$

Therefore

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) \leq \mu(\mathcal{U}_k)\mu(\mathcal{U}_n) + \psi_{\Delta}\epsilon_k\epsilon_n.$$

In Figure 2 we represented the sets $\mathcal{U}_n \cap \sigma^{-i}\mathcal{U}_n$ and $\mathcal{U}_k \cap \sigma^{-i}\mathcal{U}_n$. The light grey rectangle at the top represents the set \mathcal{U}_n and the one at the bottom the set $\sigma^{-i}\mathcal{U}_n$. The darker grey rectangle represents the set $\mathcal{U}_k \supset \mathcal{U}_n$. The “gap” Δ between the coordinates fixed by the sets \mathcal{U}_k and $\sigma^{-i}\mathcal{U}_n$ allows to use the ψ -mixing condition of the measure μ to conclude inequality (15).

Suppose now that $\eta_n < n + 1$ and that \mathcal{V}_n has coordinates fixed only in $\{-n, \dots, -n + \lfloor \eta_n/2 \rfloor\}$ (the case that \mathcal{V}_n has coordinates fixed only in $\{n - \lfloor \eta_n/2 \rfloor, \dots, n - 1\}$ is similar), then

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) \leq \mathbb{E}(\mathbb{1}_{\mathcal{V}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) = \mu(\mathcal{V}_n \cap \sigma^{-i}\mathcal{U}_n)$$

and

$$|\mu(\mathcal{V}_n \cap \sigma^{-i}\mathcal{U}_n) - \mu(\mathcal{V}_n)\mu(\mathcal{U}_n)| \leq \psi_{\lfloor \eta_n/2 \rfloor}\mu(\mathcal{V}_n)\epsilon_n. \quad (16)$$

Therefore

$$\mathbb{E}(\mathbb{1}_{\mathcal{U}_n} \cdot \mathbb{1}_{\mathcal{U}_n} \circ \sigma^i) \leq \mu(\mathcal{V}_n)\mu(\mathcal{U}_n) + \psi_{\lfloor \eta_n/2 \rfloor}\mu(\mathcal{V}_n)\epsilon_n.$$

The gap $\lfloor \eta_n/2 \rfloor$ between the coordinates fixed by the sets \mathcal{V}_n and $\sigma^{-i}\mathcal{U}_n$ allows to use the ψ -mixing condition of the measure μ to conclude inequality (16). \square

- [6] R. Bowen, *Equilibrium states and the Ergodic theory of Anosov diffeomorphisms*, (Lecture Notes in Math. 470, Springer, Berlin, 1975).
- [7] H. Bruin and S. Vaienti, Return time statistics for unimodal maps, *Fundam. Math.* **176** (2003) 77–94.
- [8] J-R. Chazottes and P. Collet, Poisson approximation for the number of visits to balls in nonuniformly hyperbolic dynamical systems, *Ergodic Theory Dynam. Systems* **33** (2013).
- [9] J-R. Chazottes and Z. Coelho and P. Collet, Poisson processes for subsystems of finite type in symbolic dynamics, *Stoch. Dyn.* **9** (2009).
- [10] Z. Coelho, Asymptotic laws for symbolic dynamical systems, *Topics in symbolic dynamics and applications (Temuco, 1997)*, *London Math. Soc. Lecture Note Ser.*, Cambridge Univ. Press, Cambridge **279** (2000).
- [11] P. Collet and J-P. Eckmann, *Concepts and results in chaotical dynamics*, (Springer, 2006).
- [12] P. Collet and A. Galves, Asymptotic distribution of entrance times for expanding maps of the interval, *Dynamical systems and applications* **WSSIAA 4** (1995) 139–152.
- [13] P. Collet and A. Galves, Statistics of close visits to the indifferent fixed point of an interval map, *J. Stat. Phys.* **72** (1993) 459–478.
- [14] W. Döebelin, Remarques sur la théorie métrique des fractions continues, *Compositio Math.* **7** (1940) 353–371.
- [15] D. Dolgopyat, Limit theorems for partially hyperbolic systems, *Trans. Amer. Math. Soc.* **356** (2004) 1637–1689.
- [16] A. Galves and B. Schmitt, Inequalities for hitting times in mixing dynamical systems, *Random Computat. Dynam.* **5** (1997) 337–348.
- [17] N.T.A. Haydn, Entry and return times distribution, *Dyn. Syst.* **28** (2013) 333–353.
- [18] N.T.A. Haydn and Y. Lacroix and S. Vaienti, Hitting and return times in ergodic dynamical systems, *Ann. Probab.* **33** (2005) 2043–2050.
- [19] N.T.A. Haydn and Y. Psiloyenis, Return times distribution for Markov towers with decay of correlations, *Nonlinearity* **27** (2014) 1323–1349.
- [20] M. Hirata, Poisson law for Axiom A diffeomorphisms, *Ergodic Theory Dynam. Systems* **13** (1993) 533–556.
- [21] M. Hirata, Poisson law for the dynamical systems with “self-mixing” conditions, *Dynamical systems and Chaos* **1** (1995) 87–96.

- [22] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, 2nd edn. (Cambridge University Press, 1995).
- [23] B. Pitskel, Poisson limit law for Markov chains, *Ergodic Theory Dynam. Systems* **11** (1991) 501–513.
- [24] Y. Kifer, Nonconventional Poisson limit theorems, *Israel J. Math.* **195** (2013) 373–392.
- [25] X. Zhang, Note on limit distribution of normalized return times and escape rate, *Stoch. Dyn.* **16** (2016).